Nonequilibrium work relations
II. Microscopic systems driven away from equilibrium

Motivating the second law from microscopic principles

for a thermally isolated system, initially at equilibrium:
$W_{cyc} \geq 0$ (Thomsom's formulation)

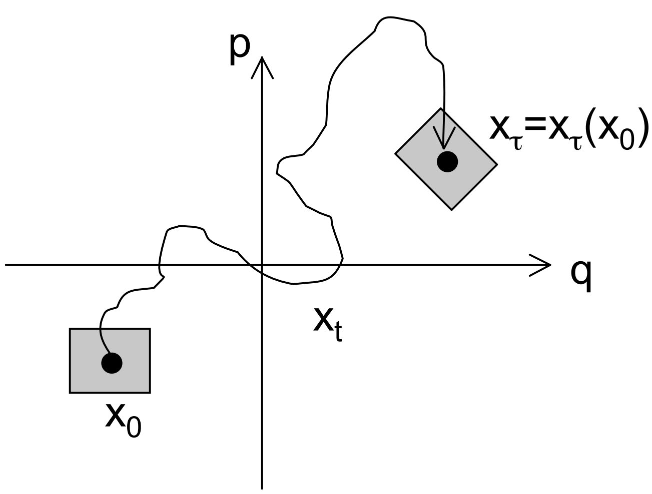
How much can we understand on the basis of microscopic (Hamiltonian) dynamics?

$x = (q,p) = (\vec{r}_1, \cdots \vec{r}_m, \vec{p}_1, \cdots \vec{p}_m)$ microstate

$H(x,\lambda)$ Hamiltonian

$\dot{q} = \frac{\partial H}{\partial p} \quad , \quad \dot{p} = -\frac{\partial H}{\partial q} \longrightarrow x(t)\,,\, x_t$ trajectory

Liouville's theorem: conservation of phase space volume



$\dot{x} = \dot{x}(x,\lambda)$

flow field

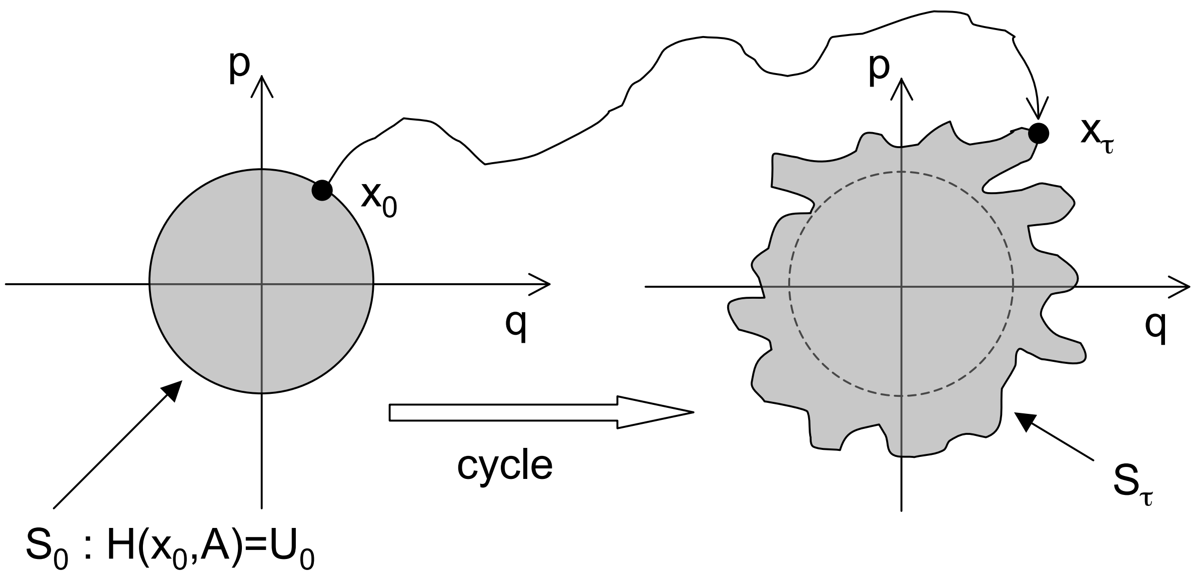
$$\nabla \cdot \dot{x} = \sum_i \left[\frac{\partial}{\partial q_i}(\dot{q}_i) + \frac{\partial}{\partial p_i}(\dot{p}_i) \right] = 0$$

$$\left| \frac{\partial x_\tau}{\partial x_0} \right| \equiv \det \left\{ \frac{\partial x_{\tau,i}}{\partial x_{0,j}} \right\} = \exp\left(\int_0^\tau \nabla \cdot \dot{x}\, dt \right) = 1$$

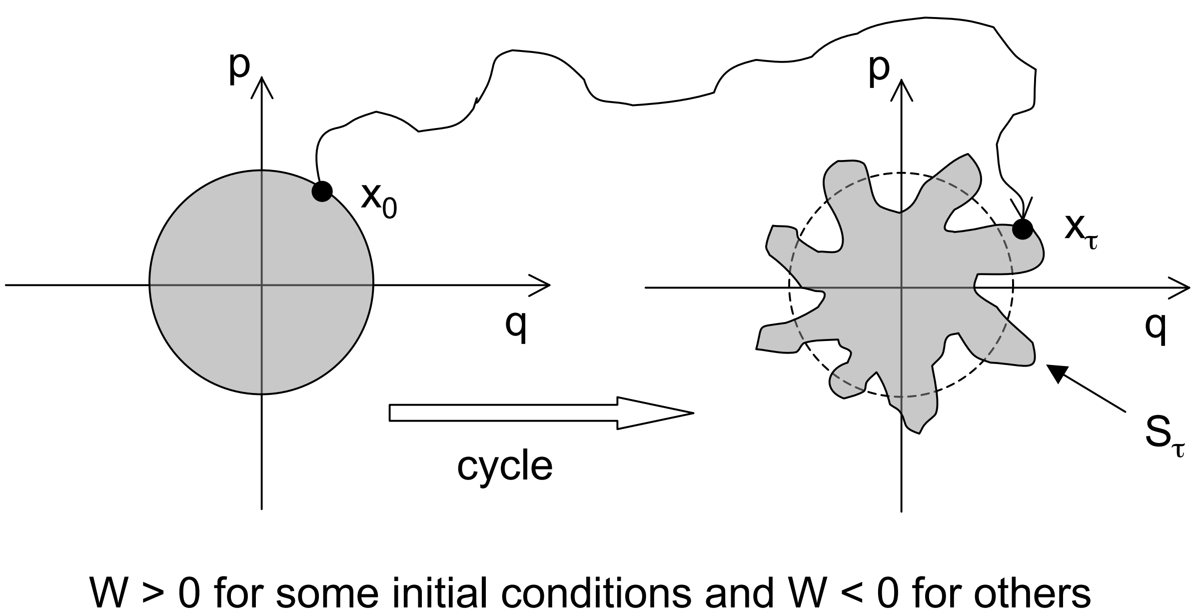
isolated system :   work = change in internal energy

$$W = H(x_\tau, B) - H(x_0, A)$$

for cyclic processes, A = B, and the 2nd law predicts W ≥ 0 …
should we expect that $H_{final} \geq H_{initial}$ for *all* initial conditions? No !



(What's wrong with this picture?)



W > 0 for some initial conditions and W < 0 for others

Interpret the 2nd law *statistically:*

$$\langle W \rangle \equiv \int dx_0\, f_A(x_0)\left[H_A\left(x_\tau(x_0)\right) - H_A(x_0)\right] \geq 0$$

where f_A is some distribution of initial conditions meant to represent the initial equilibrium state.

For what choices of f_A can we show that <W> ≥ 0 ?

Three candidates:

$$f_A^\mu(x) = \frac{1}{\Sigma}\delta\left[U_A - H_A(x)\right] \qquad \text{microcanonical}$$

$$f_A^c(x) = \frac{1}{Z}\exp\left[-\beta H_A(x)\right] \qquad \text{canonical}$$

$$f_A^M(x) = \frac{1}{\Omega}\theta\left[U_A - H_A(x)\right] \qquad \text{“macrocanonical”*}$$

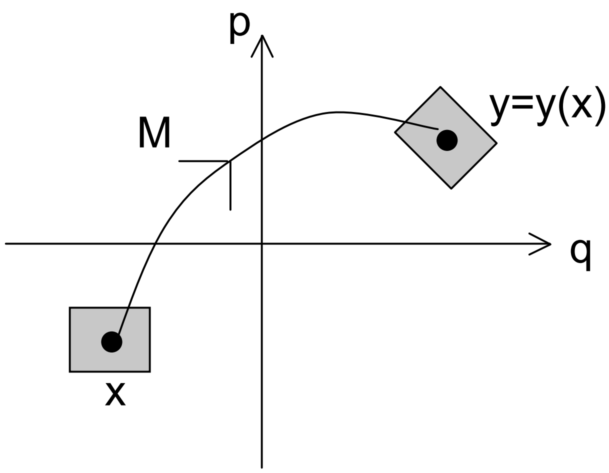
It’s easy to show that <W> ≥ 0 when initial conditions are sampled from the canonical ensemble …

• WARNING: Usually this term denotes an ensemble describing an open thermodynamic system, at fixed chemical potential(s)

Proof of Thomson’s formulation of the second law
for canonically sampled initial conditions

Let M : x→y denote an invertible, volume-preserving mapping
of phase space onto itself

[Jar02,Obe05]



$$\left|\frac{\partial y}{\partial x}\right| = 1$$

M^{-1} exists

$$\Phi(x) \equiv H_A\big(y(x)\big) - H_A(x)$$

β = some positive constant

$$\left\langle e^{-\beta\Phi}\right\rangle = \int dx\, \frac{1}{Z} e^{-\beta H_A(x)}\, e^{-\beta\left[H_A(y(x)) - H_A(x)\right]}$$

$$= \frac{1}{Z}\int dx\;\, e^{-\beta\, H_A(y(x))} = \frac{1}{Z}\int dy\;\, e^{-\beta\, H_A(y)} = 1$$

Combine w/ Jensen’s inequality: $1 = \left\langle e^{-\beta\Phi}\right\rangle \geq e^{-\beta\langle\Phi\rangle} \longrightarrow \left\langle \Phi \right\rangle \geq 0$

thus $\left\langle H_A\big(y(x)\big)\right\rangle \geq \left\langle H_A(x)\right\rangle$

Very general result: on average, the energy increases when M : x→y

Hamiltonian evolution: M : $x_0 \to x_\tau$

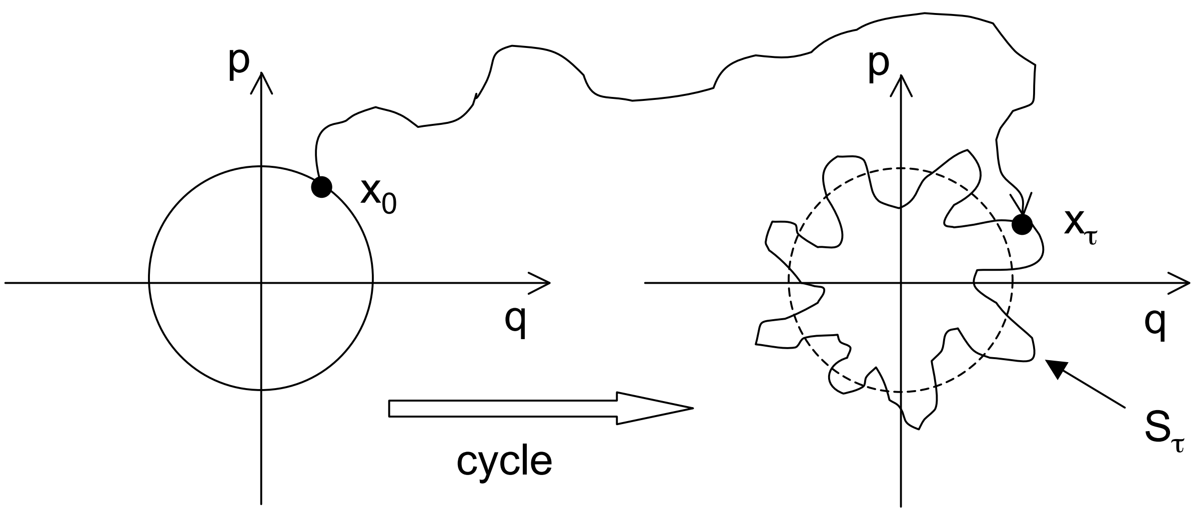
$$\therefore \left\langle H_A(x_\tau)\right\rangle \geq \left\langle H_A(x_0)\right\rangle$$

Exercise. Provide simple arguments to show that Thomson's formulation of the second law is valid when the initial conditions are sampled from the "macrocanonical" ensemble.

$$f_A^M(x) = \frac{1}{\Omega}\theta\left[U_A - H_A(x)\right]$$

What about the microcanonical ensemble?

$$f_A^\mu(x) = \frac{1}{\Sigma}\delta\left[U_A - H_A(x)\right]$$

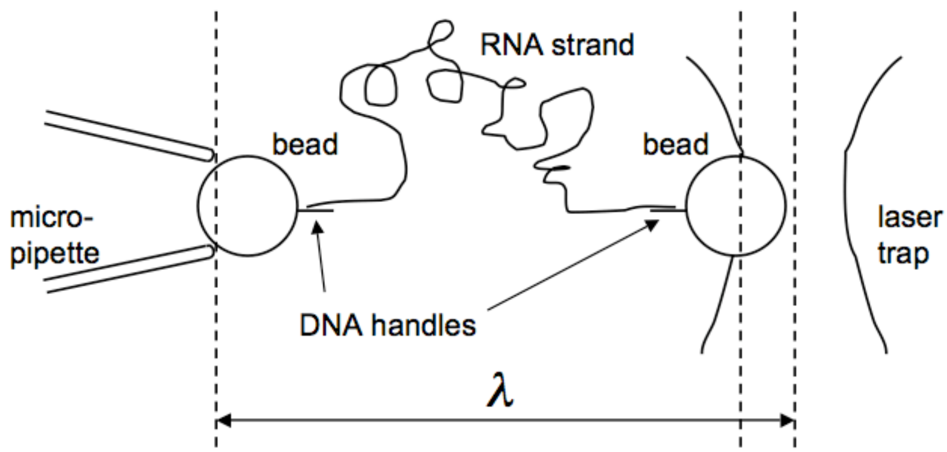


no general proof that <W> ≥ 0 in this case … for very large systems we generically expect the three ensembles will produce the same averages (*equivalence of ensembles)*

$$\langle W\rangle^\mu \approx \langle W\rangle^c \approx \langle W\rangle^M$$

Nonequilibrium work and fluctuation relations

W ≥ ΔF *on average* … what about fluctuations ?



1. Start at λ=A , in equilibrium w/ water

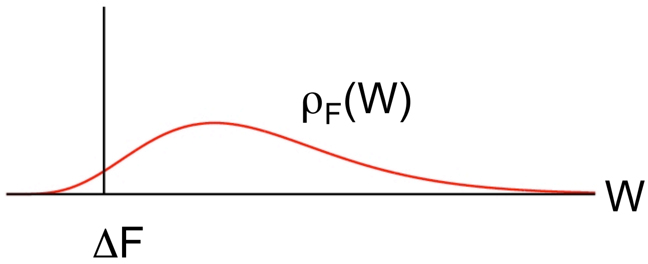
forward { 2. Stretch rapidly, λ : A→B
3. Hold λ fixed and allow to re-equilibrate

reverse { 4. "Unstretch" rapidly, λ : A←B
5. Hold λ fixed and allow to re-equilibrate

6. *Repeat*

fluctuations .. just "boring noise"? NO !!

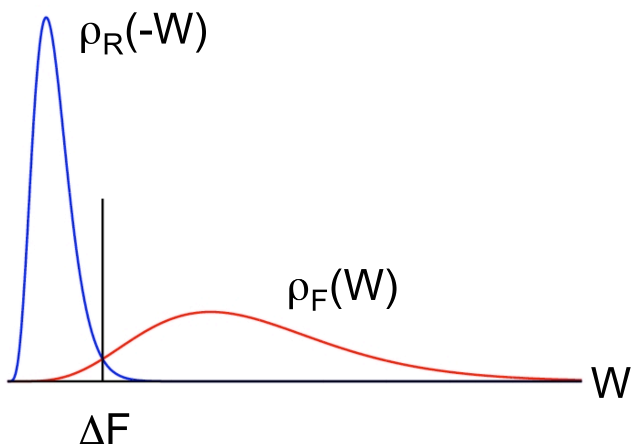
There are a number of strong and unexpected predictions we can make about the fluctuations of the work values …



$$\boxed{\left\langle e^{-\beta W} \right\rangle = e^{-\beta \Delta F}}$$

[Jar97a]

consequences: $\begin{cases} \langle W \rangle \geq \Delta F \\ P\left(W \leq \Delta F - n\beta^{-1}\right) \leq e^{-n} \end{cases}$



$$\boxed{\frac{\rho_F(+W)}{\rho_R(-W)} = e^{\beta(W-\Delta F)}}$$

[Cro98,Cro99]

consequences:

- the distributions cross exactly at W=ΔF (!)
- $\rho_F(+W)e^{-\beta W} = \rho_R(-W)e^{-\beta \Delta F}$

  integrate over W $\rightarrow$ $\left\langle e^{-\beta W} \right\rangle_F = e^{-\beta \Delta F}$

$$\boxed{\left\langle \delta(x - x_t) e^{-\beta w_t} \right\rangle = \frac{1}{Z_A} e^{-\beta H(x,\lambda_t)}}$$

[Hum01]

x = point in phase space (independent variable)
x_t = microscopic trajectory during one realization
w_t = work performed up to time t
λ_t = value of work parameter at time t
<…> = average over ensemble of realizations

l.h.s = "weighted" phase space density
r.h.s. $\propto$ equilibrium distribution

Equivalent formulation:

$$f(x,t)\left\langle e^{-\beta w_t} \right\rangle_{x,t} = f^{eq}(x,\lambda_t) e^{-\beta \Delta F(t)}$$

[Jar97b]

where $\langle \ldots \rangle_{x,t}$ denotes an average over all trajectories that pass through x at time t, f^{eq} is the canonical distribution, and $\Delta F(t) = F_{\lambda(t)} - F_A$

<u>Exercise</u>. Show that these two predictions are equivalent.

consequences:

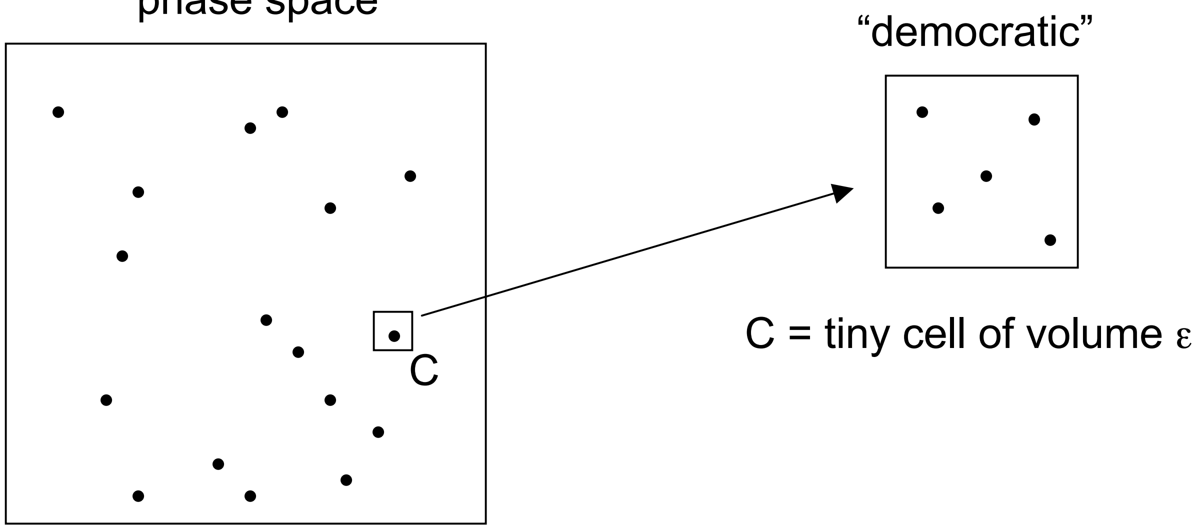
- integrate both sides (at t=τ) $\rightarrow$ $\left\langle e^{-\beta W} \right\rangle = e^{-\beta \Delta F}$
- from an ensemble of *nonequilibrium* trajectories, we can reconstruct *equilibrium* ensembles (!)

ordinarily we use the time-dependent phase space density to represent the evolution of an ensemble of trajectories:

$$f(x,t) = \langle \delta(x - x_t) \rangle$$

be careful w/ ordering of limits …

phase space



"snapshot" at time t

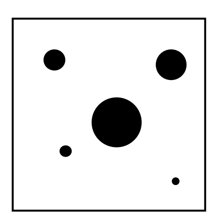
$$f(x,t) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \lim_{N \to \infty} \frac{1}{N} \sum_{x_t \in C} 1$$

another statistical representation *of the same ensemble* :

$$g(x,t) = \langle \delta(x - x_t) \exp(-\beta w_t) \rangle$$

$$= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \lim_{N \to \infty} \frac{1}{N} \sum_{x_t \in C} \exp(-\beta w_t)$$

"undemocratic"



[Jar01]

claim : $g(x,t) \propto f^{eq}(x,\lambda_t)$

$$
\begin{cases}
\textcircled{1} & \left\langle e^{-\beta W} \right\rangle = e^{-\beta \Delta F} \\[2ex]
\textcircled{2} & \dfrac{\rho_F(+W)}{\rho_R(-W)} = e^{\beta(W-\Delta F)} \\[2ex]
\textcircled{3} & \left\langle \delta(x - x_t) e^{-\beta w_t} \right\rangle = \dfrac{1}{Z_A} e^{-\beta H(x,\lambda_t)}
\end{cases}
$$

general remarks:

- exact results, valid far from equilibrium
- place strong, unexpected constraints on fluctuations
- nonequilibrium fluctuations encode equilibrium information
- closely related to other results -
    - Bochkov & Kuzovlev , $\langle \exp(-\beta W') \rangle = 1$ [Boc77,Boc81]
    - fluctuation theorems for entropy , $p(+\sigma)/p(-\sigma) = \exp(\sigma)$
    - see Udo Seifert's lectures !! [Sei05]

Derivations

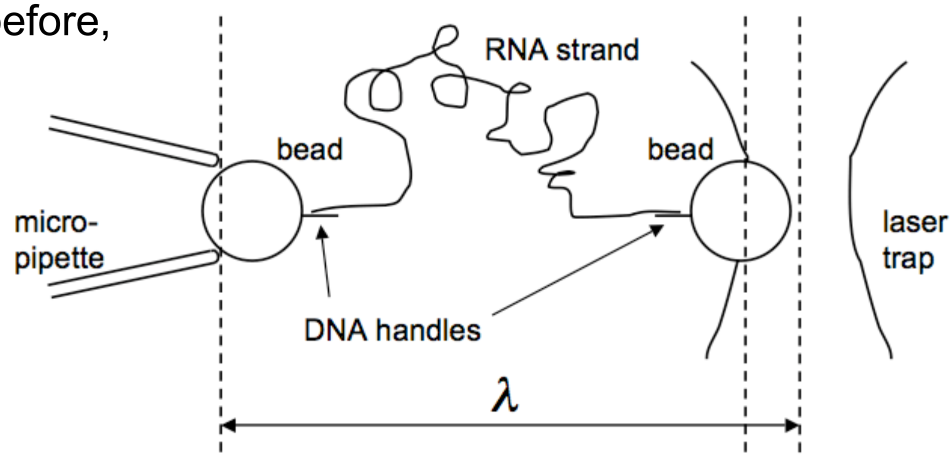
various ways to model microscopic evolution:

- Hamiltonian dynamics
- Langevin dynamics ("Brownian motion")
- discrete-time Markov processes
- deterministic thermostats (Gaussian, Nosé-Hoover)
- quantum dynamics

  & others

start w/ derivation of 1 using Hamilton's equations

as before,



$$x = (q,p) = (\vec{r}_1,\cdots\vec{r}_m,\vec{p}_1,\cdots\vec{p}_m) \qquad \text{microstate}$$

$$H(x,\lambda) \qquad \text{Hamiltonian}$$

$$\dot{q} = \frac{\partial H}{\partial p} \quad , \quad \dot{p} = -\frac{\partial H}{\partial q} \longrightarrow x(t)\,,\, x_t \qquad \text{trajectory}$$

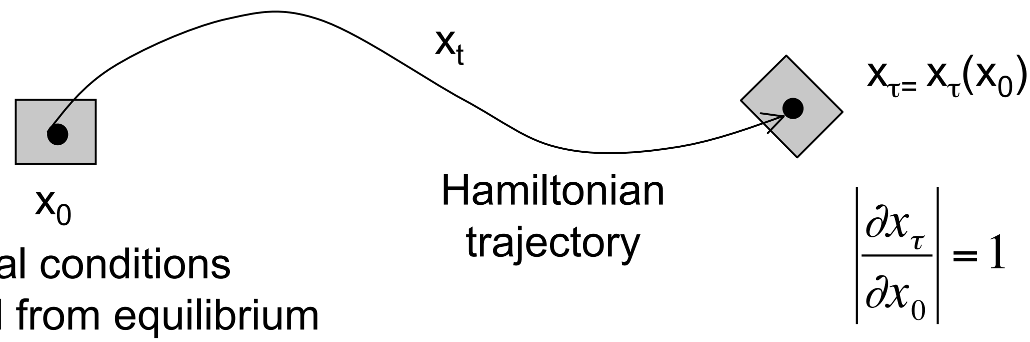
$$\lambda : A \rightarrow B \qquad \lambda(t)\,,\, \lambda_t \qquad \text{protocol}$$

For simplicity, consider *special case* of a system that is thermally isolated as the work parameter is varied, 0<t<τ
(… not plausible in real single-molecule pulling experiment !)

$$W = \int_0^\tau dt\,\dot{\lambda}\frac{\partial H}{\partial \lambda}(x_t,\lambda_t) = \int_0^\tau dt\frac{d}{dt}H(x_t,\lambda_t) = H(x_\tau,B) - H(x_0,A)$$

equilibrium state
$$\begin{cases} f_\lambda^{eq}(x) = \dfrac{1}{Z_\lambda}\exp[-\beta H(x,\lambda)] \\ Z_\lambda = \int dx \exp[-\beta H(x,\lambda)] \\ F_\lambda = -\beta^{-1}\ln Z_\lambda \end{cases}$$

(proof very similar to that of Thomson's formulation of 2nd law)



$$W = W(x_0) = H(x_\tau(x_0), B) - H(x_0, A)$$

$$\left\langle e^{-\beta W} \right\rangle = \int dx_0\, f_A^{eq}(x_0)\, e^{-\beta W(x_0)}$$

$$= \frac{1}{Z_A} \int dx_0\, e^{-\beta H_A(x_0)}\, e^{-\beta\left[H_B(x_\tau(x_0)) - H_A(x_0)\right]}$$

$$= \frac{1}{Z_A} \int dx_0\; e^{-\beta H_B(x_\tau(x_0))}$$

$$= \frac{1}{Z_A} \int dx_\tau\; e^{-\beta H_B(x_\tau)} = \frac{Z_B}{Z_A} = e^{-\beta \Delta F} \quad ☺$$

only assumptions:

1. initial conditions sampled from (canonical) equilibrium
2. isolated, classical system (Hamiltonian evolution)

This derivation can be generalized to a system in contact w/ heat reservoir, e.g. biomolecule in solution; no need to assume weak coupling, provided we properly account for solvation free energy.

[Jar04]

$$H(x, y; \lambda) = H_{sys}(x; \lambda) + H_{env}(y) + H_{\text{int}}(x, y)$$

next: derivation of (2) for discrete-time, Markov dynamics

model the evolution as a sequence of “snapshots” [Cro98]

$$x_0 \to x_1 \to x_2 \cdots \to x_t \to x_{t+1} \cdots$$

not deterministic

$P_\lambda(x \to x')$ = transition probability

$$\int dx' \, P_\lambda(x \to x') = 1$$

Markov assumption: each transition is statistically independent of the previous transition

also assume *detailed balance*

$$\frac{P_\lambda(x \to x')}{P_\lambda(x \leftarrow x')} = \exp\left[-\beta\left[H_\lambda(x') - H_\lambda(x)\right]\right]$$

interpretation:

$$\frac{1}{Z_\lambda} e^{-\beta H_\lambda(x)} P_\lambda(x \to x') = \frac{1}{Z_\lambda} e^{-\beta H_\lambda(x')} P_\lambda(x \leftarrow x')$$

l.h.s. = net rate of transitions $x \to x'$ in equilibrium
r.h.s. = net rate of transitions $x \leftarrow x'$ in equilibrium

detailed balance : the two rates are equal we can use these dynamics to model a thermodynamic process during which λ is varied with time

$$\lambda_0 = A \qquad x_0 \xrightarrow{\lambda_1} x_1 \xrightarrow{\lambda_2} x_2 \cdots x_{\tau-1} \xrightarrow{\lambda_\tau} x_\tau \qquad \lambda_\tau = B$$

(increments in λ alternate w/ Markov transitions $x_t \to x_{t+1}$)

how are *heat* and *work* represented in this scheme ?

$$Q = \sum_{t=1}^{\tau} \left[H(x_t, \lambda_t) - H(x_{t-1}, \lambda_t) \right]$$

= sum of energy changes due to Markov transitions

$$W = \sum_{t=1}^{\tau} \left[H(x_{t-1}, \lambda_t) - H(x_{t-1}, \lambda_{t-1}) \right]$$

= sum of energy changes due to changes in λ

$$\rightarrow \quad Q + W = H(x_\tau, B) - H(x_0, A) \quad \text{(First Law)}$$

consider two such processes (forward & reverse):

forward protocol $\quad A = \lambda_0^F \rightarrow \lambda_1^F \rightarrow \cdots \rightarrow \lambda_\tau^F = B$

reverse protocol $\quad A = \lambda_\tau^R \leftarrow \cdots \leftarrow \lambda_1^R \leftarrow \lambda_0^R = B$

$$\lambda_t^F = \lambda_{\tau-t}^R$$

probability to observe a particular trajectory during the forward process

$$P^F\left(x_0 \to x_1 \cdots \to x_\tau\right) = f_A^{eq}\left(x_0\right) P_{\lambda_1}\left(x_0 \to x_1\right) \cdots P_{\lambda_\tau}\left(x_{\tau-1} \to x_\tau\right)$$

sample initial conditions
from (canonical) equilibrium

transition probabilties

(similar expression for reverse process)

compact notation:

$X = x_0 \to x_1 \cdots \to x_\tau$ a forward trajectory

$X^+ = x_0 \leftarrow x_1 \cdots \leftarrow x_\tau$ its "conjugate twin"

<u>Exercise</u>. Within this framework, derive the following identity :

$$\frac{P^F\left(X\right)}{P^R\left(X^+\right)} = \exp\left[\beta\left(W^F - \Delta F\right)\right]$$

where the numerator on the left denotes the net probability to observe a particular trajectory during the forward process; the denominator is the net probability to observe its conjugate twin during the reverse process; and $W^F=W^F(X)$ is the work performed on the trajectory during the forward process.

(follow-up exercise: derive an analogous identity for the case of deterministic, Hamiltonian dynamics) finally,

$$\rho_F(W) = \int dX\, P^F(X)\,\delta\big[W - W^F(X)\big]$$

$$= \int dX\, P^R(X^+)\,e^{\beta\left[W^F(X)-\Delta F\right]}\,\delta\big[W - W^F(X)\big]$$

$$= e^{\beta(W-\Delta F)} \int dX^+\, P^R(X^+)\,\delta\big[W + W^R(X^+)\big]$$

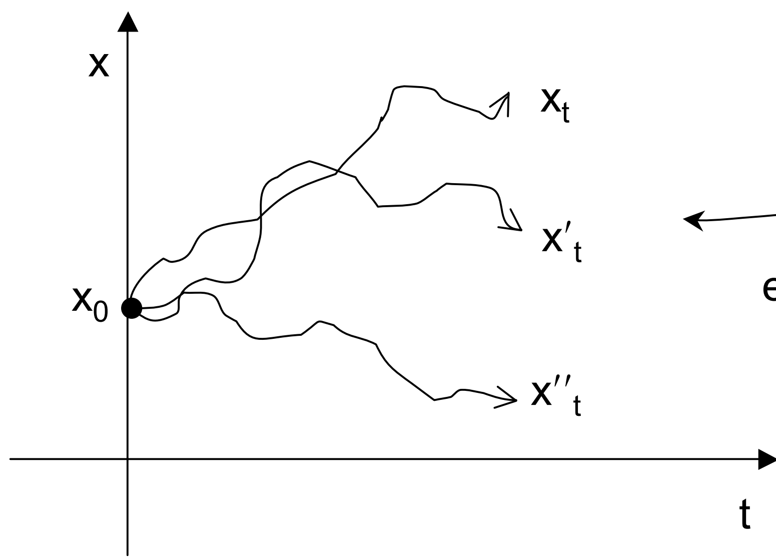
$$= e^{\beta(W-\Delta F)} \rho_R(-W)$$

$$\left(\begin{array}{c} dX = dx_0\, dx_1 \cdots dx_\tau = dX^+ \\ W^F(X) = -W^R(X^+) \end{array}\right)$$

$$\frac{\rho_F(+W)}{\rho_R(-W)} = e^{\beta(W-\Delta F)}$$

☺

Finally, let's derive 3 for continuous-time Markov trajectories
(deterministic / stochastic)



e.g. Brownian motion

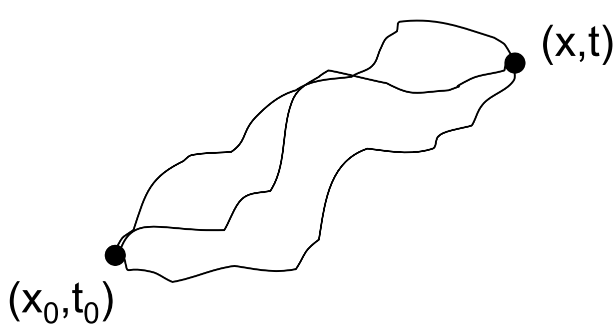
preliminaries:

$P\left[X \mid x_0, t_0\right]$ probability to observe trajectory X, given initial conditions x_0 at time t_0 (w/ respect to some measure on path space)

$K\left(x, t \mid x_0, t_0\right)$ *propagator* , probability density to arrive at x at time t, given x_0 at time t_0

$$K\left(x, t \mid x_0, t_0\right) = \int \text{đ}X\, P\left[X \mid x_0, t_0\right]$$

integral over all X with fixed initial and final conditions (x_0,t_0) & (x,t)



e.g. "sum" over these trajectories

$$\frac{\partial}{\partial t}K\left(x,t \mid x_0,t_0\right) = \int dx' R_t\left(x' \to x\right) K\left(x',t \mid x_0,t_0\right) = \hat{L}_t K$$

instantaneous transition rate

Conservation of probability: $\int dx\, R_t\left(x' \to x\right) = 0$

Feynman-Kac theorem:

Consider a function ω(x,t) and define

$$\Omega\left[X\right] = \int_{t_0}^{t} ds\, \omega\left(x_s, s\right)$$

$$G\left(x,t \mid x_0,t_0\right) = \int \mathcal{D}X\, P\left[X \mid x_0,t_0\right] e^{\Omega[X]}$$

G is like to the propagator K - it is obtained by summing over trajectories - only now each trajectory carries a statistical weight Ω

… similar to path-integral approach to quantum mechanics, only the weights exp(Ω) are real , unlike the phases $\exp\left(iS/\hbar\right)$

We have an evolution equation for K, and now we will obtain an evolution equation for G ("quick & dirty" derivation of the F-K theorem).

$$G\left(x, t+dt \mid x_0, t_0\right) = \int dx' \, G\left(x, t+dt \mid x', t\right) G\left(x', t \mid x_0, t_0\right)$$

$$\begin{aligned} G\left(x, t+dt \mid x', t\right) &= \int \text{đ}X \, P\left[X \mid x', t\right] e^{\Omega[X]} \\ &\approx e^{\omega(x',t)\,dt} K\left(x, t+dt \mid x', t\right) \\ &\approx \left(1 + \omega \, dt\right)\left[\delta\left(x - x'\right) + dt \, R_t\left(x' \to x\right)\right] \end{aligned}$$

Substitute 2nd equation into 1st, and ignore $(dt)^2$ terms …

$$G\left(x, t+dt \mid x_0, t_0\right) = \left(1 + \omega \, dt + dt \, \hat{L}_t\right) G\left(x, t \mid x_0, t_0\right)$$

$$\frac{\partial G}{\partial t} = \left(\omega + \hat{L}_t\right) G$$ Feynman-Kac formula

unweighted propagator : $$\frac{\partial K}{\partial t} = \hat{L}_t K$$

weighted propagator : $$\frac{\partial G}{\partial t} = \left(\omega + \hat{L}_t\right) G$$

statistical weight $$\exp \int^t ds \, \omega\left(x_s, s\right)$$

Now apply this result to a system driven away from equilibrium …

$$f(x,t) = \langle \delta(x - x_t) \rangle = \int dx_0\, f_A^{eq}(x_0) K(x,t \mid x_0,0)$$

$$= \int dx_0\, f_A^{eq}(x_0) \int dX\, P[X \mid x_0,0]$$

$$\frac{\partial f}{\partial t} = \hat{L}_t f$$ (e.g. Liouville, Fokker-Planck equation)

assume: $\hat{L}_t\, e^{-\beta H(x,\lambda_t)} = 0$ ("balance")

$$g(x,t) = \left\langle \delta(x - x_t) e^{-\beta w_t} \right\rangle$$

$$= \int dx_0\, f_A^{eq}(x_0) \int dX\, P[X \mid x_0,0] e^{-\beta w_t[X]}$$

$$= \int dx_0\, f_A^{eq}(x_0) G(x,t \mid x_0,0)$$

where $\Omega[X] = -\beta w_t[X] = -\beta \int_0^t ds\, \dot{\lambda} \frac{\partial H}{\partial \lambda}$

Feynman-Kac $\rightarrow$ $$\frac{\partial g}{\partial t} = \left(-\beta \dot{\lambda} \frac{\partial H}{\partial \lambda} + \hat{L}_t \right) g$$

This evolution equation for g is solved by

$$g(x,t) = \frac{1}{Z_A} \exp[-\beta H(x,\lambda_t)]$$ Verify !

$$\left\langle \delta(x - x_t) e^{-\beta w_t} \right\rangle = \frac{1}{Z_A} \exp[-\beta H(x,\lambda_t)]$$ ☺

[Hum01,Hum05]